FEBS Letters 587 (2013) 3859–3868

FEBS Letters

journal homepage: www.FEBSLetters.org

Crystal structures of the human histone H4K20 methyltransferases SUV420H1 and SUV420H2

Hong Wu [a], Alena Siarheyeva [a], Hong Zeng [a], Robert Lam [a], Aiping Dong [a], Xian-Hui Wu [a,b], Yanjun Li [a], Matthieu Schapira [a,c], Masoud Vedadi [a,c,*], Jinrong Min [a,d,*]

[a] Structural Genomics Consortium, University of Toronto, Toronto, Ontario, Canada
[b] School of Chemical Biology and Biotechnology, Laboratory of Chemical Genomics, Shenzhen Graduate School of Peking University, Shenzhen 518055, China
[c] Department of Pharmacology and Toxicology, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada
[d] Department of Physiology, University of Toronto, Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 7 July 2013
Revised 29 September 2013
Accepted 15 October 2013
Available online 26 October 2013

Edited by Christian Griesinger

Keywords:
Histone methyltransferase
SET domain
Crystal structure
SUV420H1
SUV420H2

ABSTRACT

SUV420H1 and SUV420H2 are two highly homologous enzymes that methylate lysine 20 of histone H4 (H4K20), a mark that has been implicated in transcriptional regulation. In this study, we present the high-resolution crystal structures of human SUV420H1 and SUV420H2 in complex with SAM, and report their substrate specificity. Both methyltransferases have a unique N-terminal domain and Zn-binding post-SET domain, and prefer the monomethylated histone H4K20 as a substrate in vitro. No histone H4K20 trimethylation activity was detected by our radioactivity-based assay for either enzyme, consistent with the presence of a conserved serine residue that forms a hydrogen bond with the target lysine side-chain and limits the methylation level.

1. Introduction

Histone lysine methylation plays an important role in the regulation of dynamic chromatin structure and chromatin template-mediated biological processes, including transcription, DNA repair and replication [1]. Histone lysine methylation is carried out by histone lysine methyltransferases (HMT). In the human genome, there exist about 50 histone lysine methyltransferases, and most of them contain a SET domain as their catalytic domain [2,3]. So far, six histone lysine methylation sites have been identified on histone H3 (K4, K9, K27, K36 and K79) and histone H4 (H4K20), and each of these lysine residues can be mono-, di-, or trimethylated often with distinct functional consequences. For example, SETD7 (also known as SET7/9) is a histone H3 monomethylase specifically methylating histone H3K4 [4], G9a is a histone H3K9 dimethylase [5], SUV39H1 and SUV39H2 are histone H3K9 trimethylases [5], and the three NSD family members (NSD1-3) primarily dimethylate histone H3K36 [6–8].

H4K20 is the only lysine methylation site on histone H4, and is conserved from yeast to human [9–13]. In asynchronous HeLa cells, the global levels of H4K20me1, H4K20me2, and H4K20me3 methylation is approximately 10%, 80%, and 5%, respectively [14]. Methyltransferases that specifically methylate H4K20 include the monomethylase SETD8 (or PR-Set7) [10–12], SUV420H1 and SUV420H2, which are primarily H4K20 dimethylases [14–17]. Although NSD1, NSD2 and ASH1 have also been reported to be able to methylate H4K20 in vitro [18–20], depletion of ASH1 or NSD1 did not affect global levels of K20H4me1, me2 or me3 [21]. Also Gozani's lab were not able to observe methylation activity of NSD2 at H4K20 despite using multiple approaches [8].

The H4K20me1, H4K20me2, and H4K20me3 marks have been linked to transcriptional repression and activation [22–27], DNA repair [13,16,28–30], and pericentric heterochromatin formation [15], respectively. L3MBTL1, L3MBTL2 and 53BP1 have all been found to recognize H4K20me1 and H4K20me2 marks [22,23,30–35]. L3MBTL1 is a MBT repeat containing protein, and its MBT domain was found to compact chromatin dependent on mono- and di-methylation of H4K20 and H1BK26 [22]. In another report, the association of L3MBTL1 with chromatin was shown to be specifically dependent on the presence of H4K20me1, which is deposited by the H4K20 monomethylase SETD8, and this association promotes transcriptional repression [23]. It was also reported that

* Corresponding authors at: Structural Genomics Consortium, University of Toronto, Toronto, Ontario, Canada.
E-mail addresses: m.vedadi@utoronto.ca (M. Vedadi), jr.min@utoronto.ca (J. Min).

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2013.10.020

human SUV420H1 and SUV420H2 preferentially dimethylate H4K20 in vivo [14]. A similar result was observed for the product of the *Drosophila* single *Suv420* gene [36]. While it remains obscure whether the histone H4K20 methyltransferases SUV420H1 and SUV420H2 specifically generate H4K20me2 in the 53BP1-mediated DNA damage pathway, 53BP1 in mammals and Crb2 in fission yeast were demonstrated to interact with H4K20me2 through their tandem Tudor domain, and promote DNA damage checkpoint signaling at DNA double-strand break sites [13,21,28,30].

Suv420h1/h2-double-null mice are perinatally lethal; their chromatin has virtually lost all H4K20me2 and H4K20me3 marks and displays increased DNA damage repair defects, suggesting that the SUV420H1 and SUV420H2 enzymes play an essential role in genome integrity [16]. On the basis of gene knock-down analysis, it was suggested that the H4K20me1 mark created by the monomethylase SETD8 is the predominant and preferential substrate for SUV420H1 and SUV420H2 in vivo, although SUV420H1 and SUV420H2 could also utilize unmodified H4K20 as a substrate in vitro and in vivo [15,21].

In this study we show that SUV420H1 and SUV420H2 can methylate both unmodified and monomethylated H4K20 peptides with higher catalytic efficiency (k_{cat}/K_m) for monomethylated peptides. Both enzymes were active with nucleosome as a substrate but exhibited no detectable H4K20 trimethylation activity. Crystal structures of SUV420H1 and SUV420H2 provided mechanistic insights into their distinct substrate specificity and their inability in histone H4K20 trimethylation. We also report on assay optimization for screening these proteins in a 384-well format, enabling the scientific community to search for potent and selective inhibitors (chemical probes) to further investigate the role of these proteins in cells. Discovery of such compounds could provide an opportunity to control the level of activities of these proteins in cells and further investigate their involvement in cancers.

2. Materials and methods

2.1. Protein expression and purification

DNA fragments encoding the catalytic domains of human SUV420H1 (residues 63–335) and SUV420H2 (residues 2–248) was sub-cloned into pET28-MHL vector (GenBank Deposition ID: EF456735) [37], downstream of the poly-histidine coding region. The proteins were over expressed in *Escherichia coli* BL21 (DE3) codon plus RIL strain (Stratagene) by addition of 1 mM isopropyl-1-thio-D-galactopyranoside and incubated overnight at 15 °C. Harvested cells were resuspended in 50 mM HEPES buffer, pH 7.4, supplemented with 250 mM NaCl, 5 mM imidazole, 2 mM β-mercaptoethanol, 5% glycerol. The cells were lysed by passing through a microfluidizer (Microfluidics Corporation) at 20,000 psi. After clarification of the crude extract by high-speed centrifugation, the lysate was loaded onto a 5 mL HiTrap Chelating column (GE Healthcare), charged with Ni^{2+}. The column was washed with 10 column volumes of 20 mM HEPES buffer, pH 7.4, containing 250 mM NaCl, 50 mM imidazole and 5% glycerol, the protein was eluted with 20 mM HEPES buffer, pH 7.4, 250 mM NaCl, 250 mM imidazole, 5% glycerol. The proteins were loaded onto a Superdex 200 column (26 × 60) (GE Healthcare), equilibrated with 20 mM PIPES buffer, pH 6.5, and 250 mM NaCl. TEV protease was added to combined fractions containing SUV420H1 and SUV420H2 to remove His-tag. The proteins were further purified to homogeneity by ion-exchange chromatography. The mutants of SUV420H1 (residues 63–335) were generated by using QuikChange Site-direct Mutagenesis kit (Agilent Technologies). All the mutants were expressed and purified as described above.

2.2. Crystallization

Purified SUV420H1 and SUV420H2 proteins (10 mg/mL) were mixed with S-adenosyl-L-methionine (SAM) (Sigma) at 1:10 molar ratio of protein:SAM and crystallized using the sitting drop vapor diffusion method by mixing 1 μL of protein solution with 1 μL of the reservoir solution containing 2.0 M sodium formate, 0.1 M Bis-Tris propane, pH 7.0 (for SUV420H1), or 12% PEG 8000, 0.1 M HEPES pH 7.2, 12% ethylene glycol (for SUV420H2).

2.3. Data collection, structure determination and refinement

X-ray diffraction data for selenium crystal of SUV420H1 were collected at 100 K at the peak wavelength of the selenium K absorption edge (0.97904 Å) using synchrotron radiation at the Canadian Macromolecular Crystallography Facility (CMCF) on beamline 08ID-1 at the Canadian Light Source (Saskatoon, SK, Canada), while data for native SUV420H2 crystal were collected in-house using Rigaku FRE High Brilliance X-ray Generator with R-AXIS IV detector. Intensities were integrated and scaled using HKL2000 [38]. The structure of SUV420H1 was determined using the single-wavelength anomalous dispersion (SAD) method utilizing the anomalous signal from selenium atoms present in the selenomethionine-labeled protein crystal. The positions of 10 selenium anomalous scatters were found with the program SHELXD [39] followed by heavy atom refinement and maximum-likelihood-based phasing as implemented in the autoSHARP program suite [40]. Phase improvement by density modification generated an interpretable experimental SAD map, which allowed an initial model to be built using the autotracing algorithm in the ARP/wARP program [41]. Following several alternate cycles of restrained refinement against a maximum-likelihood target and manual rebuilding using COOT [42], the improved model revealed clear electron densities allowing placement of the bound ligand, S-adenosyl-methionine (SAM), and ordered solvent molecules. All refinement steps were performed using REFMAC [43] in the CCP4 program suite [44]. During the final cycles of model building, translation-libration-screw (TLS) parameterization [45] was included in the refinement of the final model which comprised protein, SAM, Zn ions, and solvent molecules. The model of SUV420H1 was then used as search model for solving the structure of SUV420H2 by molecular replacement using Molrep [46]. Data collection and refinement statistics are summarized in Table 2.

2.4. HMT enzyme assays

The peptides H4(1–24) (SGRGKGGKGLGKGGAKRHRKVLRD-K), H4(1–24)me1, H4(1–24)me2 and H4(1–24)me3 were purchased from Tufts University Core Facility. A lysine residue was added to the C-terminal of the peptide for biotinylation. ^{3}H-SAM and 96-well Flash plate coated with streptavidin were obtained from Perkin Elmer. SAM2® Biotin Capture Membrane was obtained from Promega (Cat. #V2861). Ecolite(+) Liquid Scintillation Cocktail (product #188247505) was from MP Biomedicals. Methyltransferase activities of wild-type SUV420H1 and SUV420H2 and four mutants (SUV420H1: Y306F, C305F, M253A, and F311A) were measured using a radioactivity-based assay. In this assay ^{3}H-SAM was used as a methyl donor to methylate peptide substrates. Peptide substrates were biotinylated to be captured through their interaction with streptavidin using streptavidin-coated SAM2® Biotin Capture membranes. The assay conditions were optimized for the wild-type and each mutant, and all experiments were performed at linear initial velocities. The optimal concentrations for all the SUV420H1 and SUV420H2 enzymes, with exception of SUV420H1_C305F mutant, were 0.5 μM. For SUV420H1_C305F

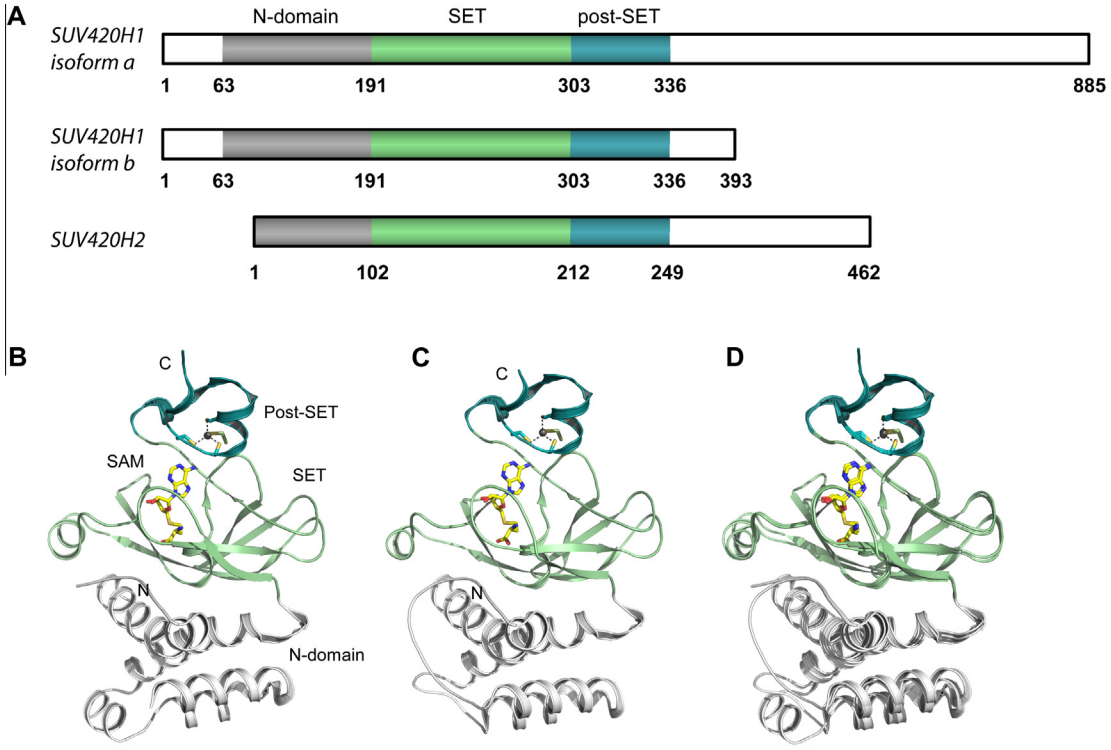

Fig. 1. Overall structure of SUV420H1 and SUV420H2. (A) Schematic illustration of the domain structures of SUV420H1 and SUV420H2. In human cells, the SUV420H1 gene has two isoforms and the two isoforms share the N-terminal SET domain. (B) Overall structure of SUV420H1. The N-terminal domain, SET domain and post-SET domain are colored in grey, green and cyan, respectively. The cofactor SAM is shown in a stick model. The four cysteines bound to the Zn ion are also shown in sticks. (C) Overall structure of SUV420H2 drawn in the same orientation as in B. (D) Superposition of SUV420H1 and SUV420H2 structures.

Table 1
Kinetic measurements of the enzymatic activities of SUV420H1, its mutants and SUV420H2.

Enzyme	Substrate	K_m^{App} (μM)		k_{cat} (h^{-1})	k_{cat}/K_m (μM^{-1} h^{-1})
		Peptide/Nucleosome	SAM		
SUV420H1	$H4_{1-24}$K20me0	3.4 ± 0.4	45 ± 6	2.8 ± 0.1	0.82
	$H4_{1-24}$K20me1	1.7 ± 0.3	9.4 ± 2.2	3.8 ± 0.3	2.24
	$H4_{1-24}$K20me2	No detectable activity			
	Nucleosome	0.3 ± 0.1	31 ± 6	1.2 ± 0.2	4
SUV420H2	$H4_{1-24}$K20me0	21 ± 2	39 ± 6	1.7 ± 0.1	0.08
	$H4_{1-24}$K20me1	16 ± 2.8	43 ± 6	4.3 ± 0.3	0.27
	$H4_{1-24}$K20me2	No detectable activity			
	Nucleosome	0.5 ± 0.1	9.5 ± 1.2	5.3 ± 0.1	10.6
SUV420H1-C305F	$H4_{1-24}$K20me1	20 ± 1.5	83 ± 5	9.6 ± 0.3	0.48
SUV420H1-Y306F		9.7 ± 1.6	53 ± 4	7.0 ± 0.3	0.72
SUV420H1-F311A		8.5 ± 1.7	80 ± 12	7.0 ± 0.7	0.82
SUV420H1-G229F		17 ± 1.6	19 ± 3	4.6 ± 0.3	0.27
SUV420H1-G229Y		No detectable activity			
SUV420H1-W264A		No detectable activity			
SUV420H1-S251A		No detectable activity			
SUV420H1-F281A		No detectable activity			

Note: The kinetic parameters presented in this table are determined under conditions described in Experimental procedures and are considered apparent values (K_m^{App}).

the enzyme concentration for all the assays was set at 2 μM. The reaction was prepared in the final volume of 30 μL. The reaction mix contained 24 μL of buffer (20 mM Tris–HCl, 10 mM DTT, pH 8.0), and 3 μL of 66 μM ^{3}H-SAM (diluted with cold SAM to achieve the desired concentrations in each assay). The reaction was started by adding 3 μL of a respective peptide substrate. The reaction mixtures were incubated for 30 min and were quenched with 30 μL of 7.5 M Guanidine-HCl. In case of peptide substrate, 7 μL of the reaction mix containing Guanidine-HCl was transferred into a SAM2® Biotin capture membrane. Membranes were washed twice with 5 mM NaCl and three times with deionized water. Membranes were dried for 2 h in the open air. Scintillation cocktail was added into each individual vial before reading the CPM (counts per minutes). Amount of the product was quantified by tracing the radioactivity (CPM) measured by a Liquid Scintillation Counter 3385 from Packard Instruments). The HMT activity was then calculated based on the amount of methylated peptide/nucleosome converted by the HMTs.

To determine the K_m values for SAM and peptide, SAM or peptide was titrated into the reaction mix that contained a saturating

Table 2
Crystallography data and refinement statistics.

	SUV420H1	SUV420H2
PDB Code	3S8P	3RQ4
Data collection		
Space group	*P*1	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	46.4, 50.1, 74.8	34.7, 60.2, 126.6
α, β, γ (°)	101.0, 108.1, 89.8	90, 90, 90
Resolution (Å) (highest resolution shell)	50.00–1.85 (1.89–1.85)	50.00–1.80 (1.85–1.80)
R_{merge}	6.8 (50.9)	11.9 (35.8)
I/sI	32.8 (2.6)	26.6 (2.5)
Completeness(%)	97.8 (96.8)	98.7 (83.3)
Redundancy	3.9 (3.9)	8.3 (3.4)
Refinement		
Resolution (Å)	33.90–1.85	50–1.80
No. reflections	52,470	24,269
R_{work}/R_{free}	18.7 / 21.2	0.178/0.221
No. atoms		
Protein	1,876	1,943
Cofactor	54	15
Zn	2	1
Water	270	232
B-factors (Å^2)		
Protein	28.5	18.8
Cofactor	23.5	12.2
Zn	27.5	15.0
Water	36.6	30.4
RMSD		
Bond lengths (Å)	0.010	0.014
Bond angles (°)	1.2	1.4
Ramachandran plot % residues		
Favored	90.0	91.4
Additional allowed	10.0	8.7
Generously allowed	0.0	0.0
Disallowed	0.0	0.0

concentration of the second substrate (peptide or SAM). The saturation of the second substrate was reached by fixing its concentration at 5 times of its respective K_m value. The K_m values were calculated using a one site saturation equation using Sigma Plot software (Systat Software, Inc., CA). Standard deviations were calculated from three independent experiments. To calculate the k_{cat} values for all the enzymes, standard curve was generated using labeled methylated H3 (1–25). To obtain labeled methylated H3 (1–25), 2 μM of H3 (1–25) substrate was incubated with 2 μM hot SAM and 200 nM G9a for 3 h to ensure that every peptide is methylated. The reaction mix was quenched as described above and standard curve was generated by using a serial dilution of the G9a/Peptide/SAM mixture and reading the CPM. All the experiments were performed in triplicates. Enzyme assays with nucleosome as a substrate was performed using a filter-based assay as previously described in Yu et al. [47].

3. Results and discussion

3.1. Enzymatic activity of SUV420H1 and SUV420H2

The SUV420 family of histone H4K20 methyltransferases was first isolated and characterized by Jenuwein's lab in 2004 [15]. There are two *Suv420* genes in vertebrates, i.e., *Suv420h1* and *Suv420h2* (Fig. 1A). In human cells, SUV420H1 exists as two alternatively spliced isoforms, which share the N-terminal SET-domain containing catalytic domain and function in substantially different chromatin locations (Fig. 1A) [48]. Although SUV420H1 and SUV420H2 preferentially dimethylate H4K20, they also mono- and trimethylate this residue [15,21,49]. These two proteins have been reported to dimethylate about 80 and 90% of total histone H4 in HeLa and S2 cells, respectively. However, they show trimethylation activity in specific sites in chromatin [21,48]. Highest H4K20 methylation level has been reported during mitosis that has been proposed to enable the transcriptionally silent state of genes [11,14,50]. At the same time, trimethylation of H4K20 in mammalian tissues has been reported to be associated with aging [51]. In order to characterize the enzymatic activity of these two H4K20 methyltransferases quantitatively, we cloned and purified the catalytic domains of both SUV420H1 (residues 63–335) and SUV420H2 (residues 2–248), which yielded well-behaved proteins for enzymatic characterization and crystallographic studies (Supplementary Fig. 1).

We developed and employed a radioactivity-based assay to evaluate the substrate specificity of these two proteins as described in Materials and methods. In this assay we monitored transfer of the methyl group from ^{3}H-SAM to peptide substrates. Both enzymes are most active at pH 8.0 and in assay conditions with no NaCl (Supplementary Fig. 2). They also tolerate well the presence of up to 5% DMSO (<15% decrease in activity) and 0.05% triton (no effect). Using the optimized conditions, we determined the kinetic parameters for both SUV420H1 and SUV420H2 activity with peptides corresponding to residues 1 to 24 of histone H4 with different H4K20 methylation states (Table 1).

Our data demonstrate that SUV420H1 and SUV420H2 mono- and di-methylate H4K20 but exhibit no detectable H4K20 trimethylation activity (Table 1, Supplementary Fig. 3). Although there is no dramatic difference in the K_m and k_{cat} values between unmethylated and monomethylated peptide substrates for either protein, both SUV420H1 and SUV420H2 showed about 3 times higher catalytic efficiency (k_{cat}/K_m) for monomethylated H4K20 substrate. This is consistent with previous reports that SUV420H1 and SUV420H2 preferentially dimethylate histone H4K20, and form only small amounts of H4K20me3 in vivo [15,21]. Considering low k_{cat} values even with dimethylated H4K20 as a substrate, trimethylation activity of these two enzymes may be below the detection limit of our assay. Presence of other interacting proteins in cells may also be required to further activate trimethylation function of these two proteins [15]. Both proteins were also active with the recombinant nucleosome as a substrate with k_{cat} values of $1.2 \pm 0.2\ h^{-1}$ and $5.3 \pm 0.1\ h^{-1}$ for SUV420H1 and SUV420H2, respectively (Table 1 and Supplementary Fig. 4).

3.2. Overall structures of SUV420H1 and SUV420H2

In order to understand the substrate specificity of SUV420H1 and SUV420H2, we determined the crystal structures of the catalytic domains of human SUV420H1 and SUV420H2 in complex with the cofactor SAM, using the same constructs we used for enzymatic assays (Fig. 1B–D). The crystal structure of SUV420H1 was solved at 1.85 Å resolution by single-wavelength anomalous dispersion (SAD) method using a seleno-methionine substituted SUV420H1 crystal (Fig. 1B). The crystal structure of SUV420H2 was solved by molecular replacement to 1.8 Å resolution (Fig. 1C). Detailed crystal diffraction data and refinement statistics are shown in Table 2.

Given that the catalytic domains of SUV420H1 and SUV420H2 share 65% sequence identity (Fig. 2A), it is not surprising that their structures are also highly conserved with an RMSD of 0.5 Å over all the aligned Cα atoms (Fig. 1D). The overall structures of the catalytic domains of SUV420H1 and SUV420H2 consist of three subdomains (Fig. 1): an N-terminal helical domain (referred to as N-domain thereafter), a SET domain and a Zn-binding post-SET domain. Compared to SUV420H2, SUV420H1 has a bigger N-domain as shown by a 22-residue insertion between the first and second helices; the majority of these residues were not observed in the crystal structure due to crystallographic disorder (Fig. 2A).

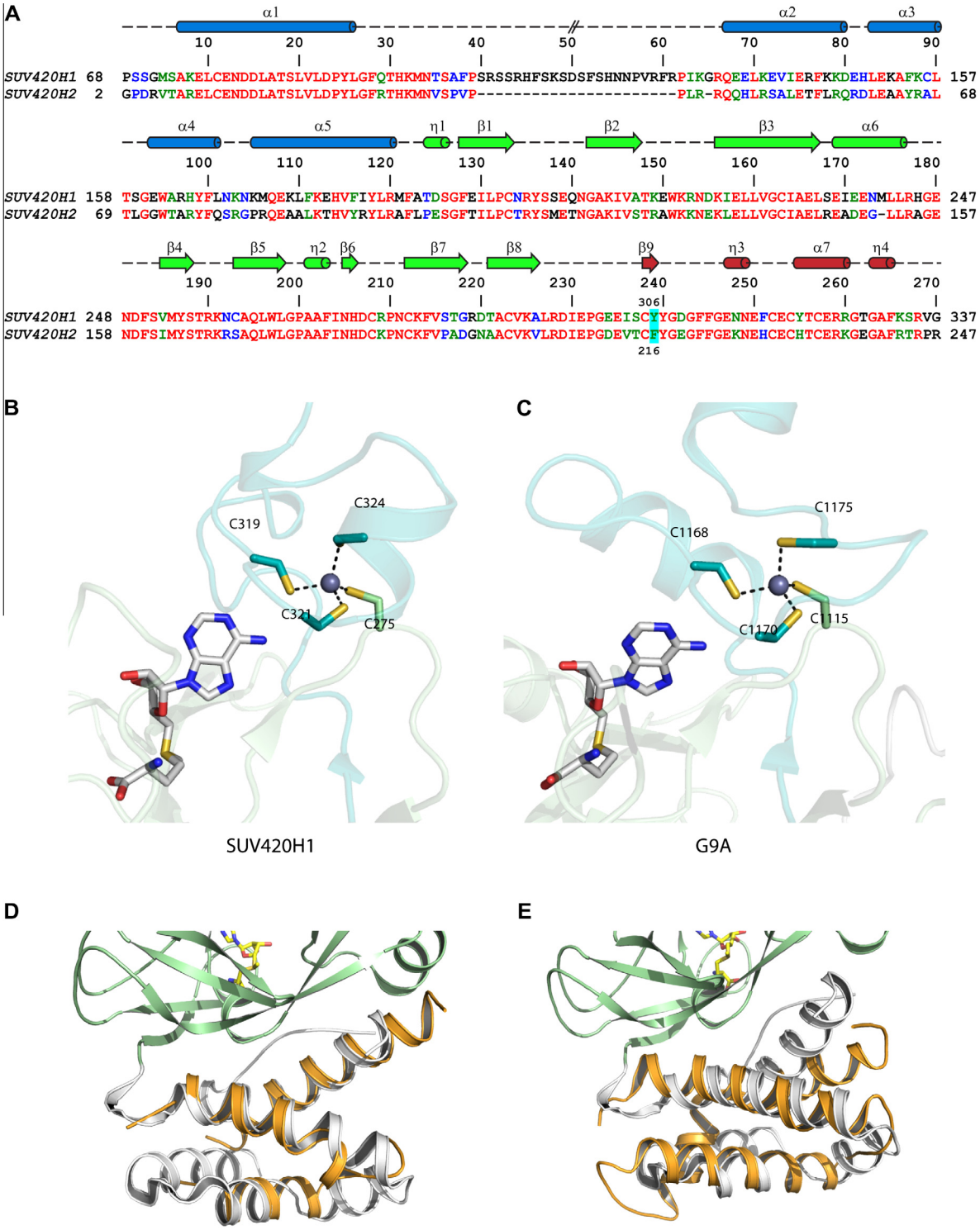

Fig. 2. (A) Structure-based sequence alignment of SUV420H1 and SUV420H2. The secondary structure elements of SUV420H1 are shown on top of the sequence, with cylinder represent helixes and arrows represent strands. The identical residues are colored in red, very similar in green, similar in blue and the rest are in black. Residues Y306 (SUV420H1) and F216 (SUV420H2) are highlighted. (B) and (C) Comparison of SUV420H1 (PDB 3S8P) and G9a (PDB 2O8J) Zn-binding domain. The Zn ion is shown in sphere, while the cysteines are shown in sticks with the three from the post-SET domain colored in cyan and the one from the SET domain in green. D and E. Comparison of the N-domain of SUV420H1 (grey) with the protein phosphatase PP2C (left, orange. PDB: 1A6Q), and T4 RNA ligase (right, orange. PDB: 2HVR).

Although the SET domains of SUV420H1 and SUV420H2 share a low sequence identity with other solved SET domain methyltransferases (identity ranging from 8% to 22%), they are highly conserved in structure. The SET and post-SET domains form a pseudo-knot, similar to other SET domain structures [52]. The SET domains harbor the active site, and both the SET and post-SET domains are involved in substrate and cofactor binding (Fig. 1) [52].

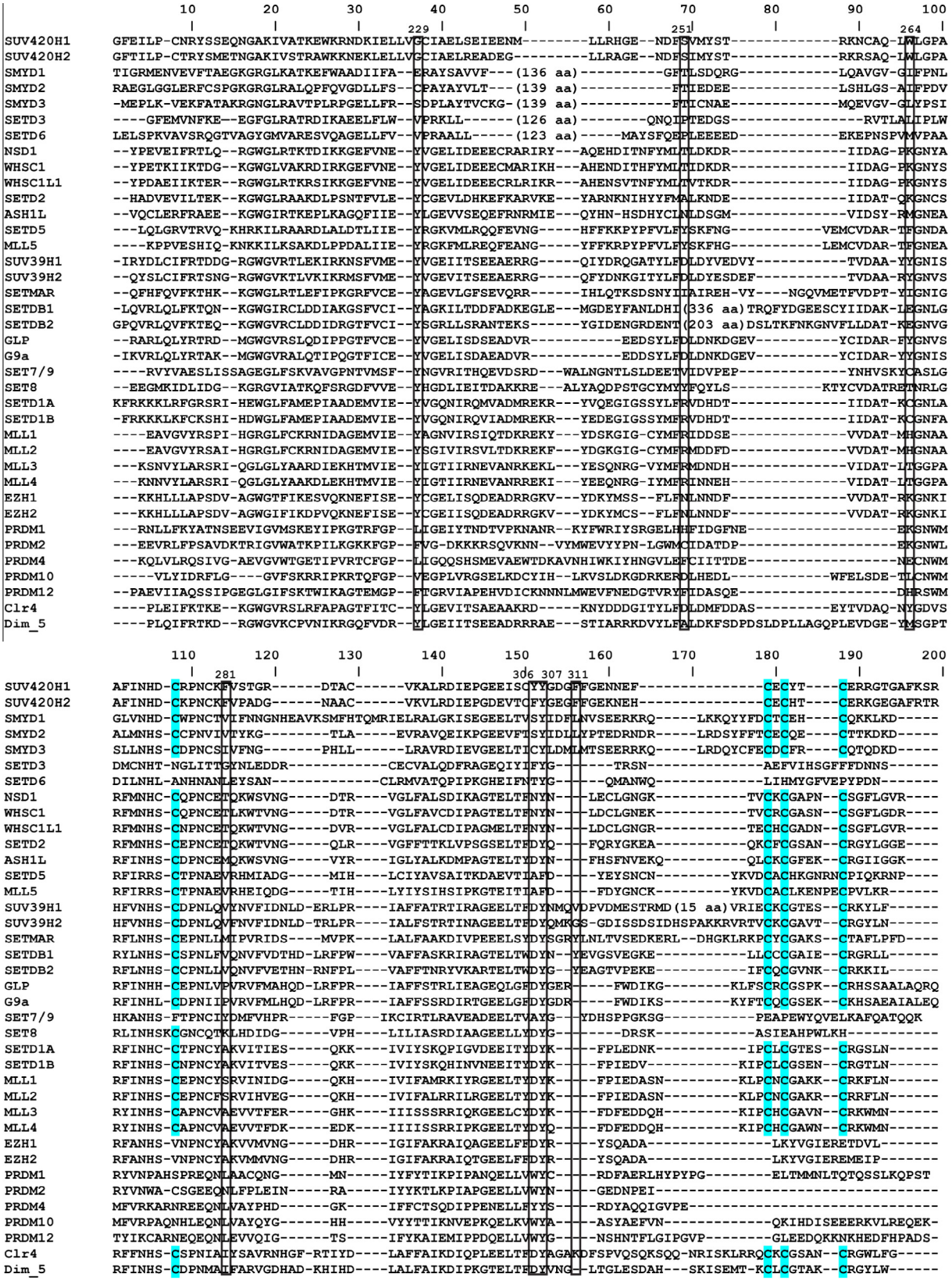

Fig. 3. Multiple sequence alignment of SET domains from different human histone methyltransferases. The cysteines bound to Zn ion are highlighted in cyan. The active site residues are boxed.

3.3. SUV420H1 and SUV420H2 have unique Zn-binding post-SET domain and N-domain

So far, over 20 SET domain methyltransferase structures have been solved. Like other Zn-binding post-SET domains, the Zn ion in SUV420H1 and SUV420H2 is bound by four cysteine residues, i.e., three cysteine residues from the post-SET domain and a fourth cysteine from the knot-like structure of the SET domain close to the SET domain active site [53]. On the basis of structural comparison and sequence alignment, we found that SUV420H1 and SUV420H2 have a unique Zn-binding domain (Fig. 2B, C and Fig 3). The three cysteine residues from the post-SET domains of SUV420H1 and SUV420H2 have a CxCxxC motif (x represents any residue), while the other SET domains have a CxCxxxxC motif (Fig. 3). Structural comparison shows that the third cysteine residue is located in a α helix in the post-SET domains of SUV420H1 and SUV420H2, whereas the third cysteine residue in other post-SET domains is located in a circular loop, which requires more residues to loop back the third cysteine residue to coordinate the Zn ion (Fig. 2B and C).

SUV420H1 and SUV420H2 have a helix bundle as their N-domain, which consists of 5 (SUV420H2) or 6 (SUV420H1) helices (Fig. 2D and E). This N-domain is unique compared to other solved SET domain methyltransferase structures. A search of the protein structure databank using the DALI server revealed that the protein serine/threonine phosphatase PP2C and the RNA ligase Rnl2 have a similar domain [54,55] (Fig. 2D and E). In PP2C, this domain is located far away from the catalytic site and is not directly involved in catalysis, although it may play some role in substrate specificity [54]. In Rnl2, the helix bundle domain is also away from the active site, but it is crucial for activity [55]. On the basis of the solved SET domain methyltransferase structures, the N-domain seems to be very diverse, and appears to be a unique and defining feature of histone methyltransferases, which prompts us to propose that incorporation of structurally divergent domains to the N terminus of the catalytic SET domain may confer different biochemical specificity to the various members of histone methyltransferases. This hypothesis requires further investigation in the future.

3.4. Active sites of SUV420H1 and SUV420H2

Based on available structures, the substrate lysine of SET-domain methyltransferases is generally surrounded with one or more polar side-chains that restrain the rotational freedom of the methyl-accepting nitrogen, which is necessary to align its lone pair with the scissile bond of the cofactor's sulfonium group. Rotational freedom–and therefore methylation level–is inversely correlated with the number of polar side-chains surrounding the substrate lysine. For example, mutating Y245 or Y305 in SETD7 to alanine changes SETD7 from a monomethylase to trimethylase [4,53]. Mutating F281 to tyrosine in Dim-5 converts Dim-5 to only

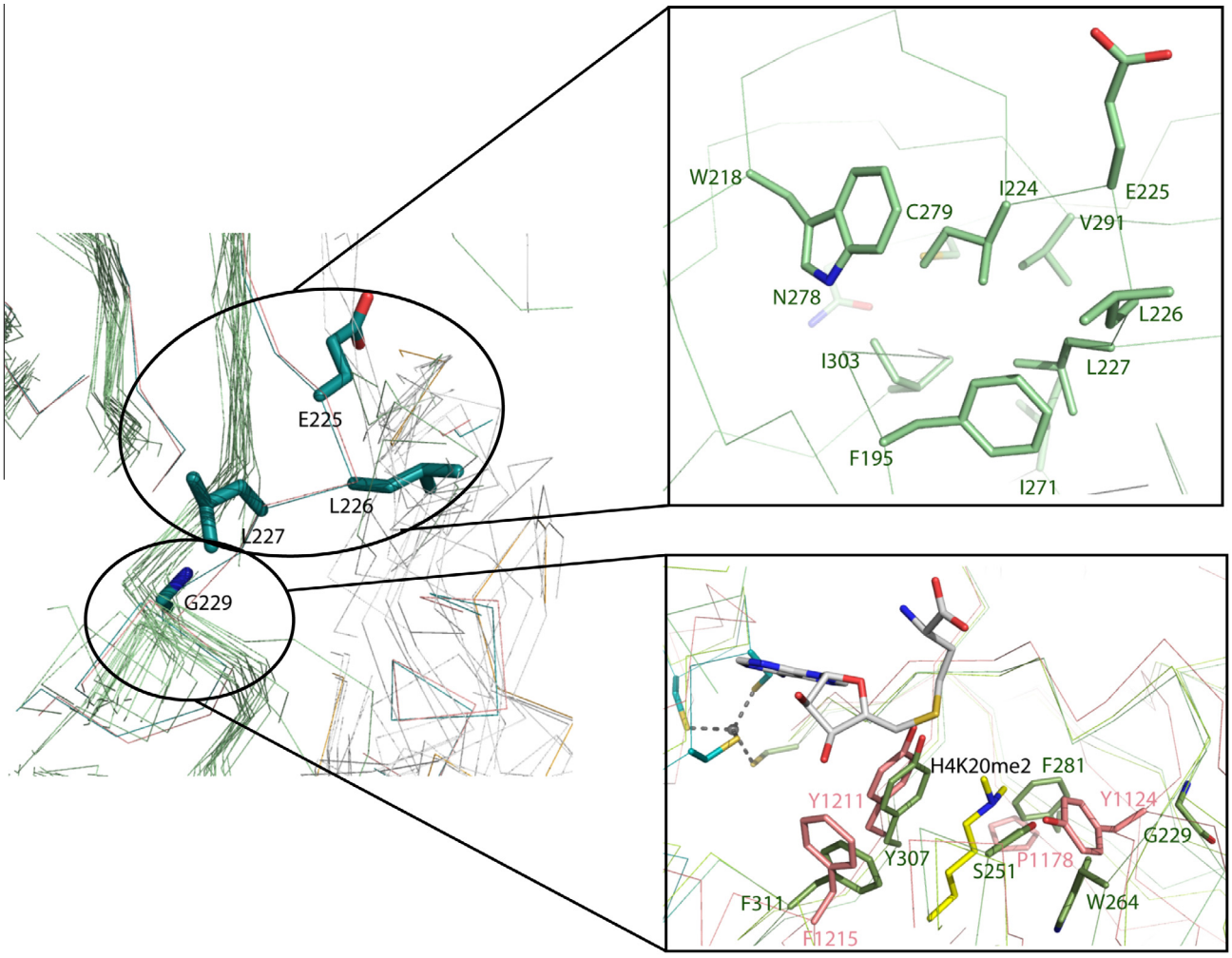

Fig. 4. Active site of SUV420H1. SUV420H1 and SUV420H2 have a unique structural bulge near the active site. A close-up view of the bulge is shown in the top right panel. Like other SET domain proteins, SUV420H1 and SUV420H2 also have an aromatic pocket, which presumably would accommodate the target lysine K20 of H4. Comparison of the active site of SUV420H1 (green) and GLP (pink, PDB 2RFI) is shown in the bottom right panel. The di-methyllysine from mouse SUV420H2 structure (PDB 4AU7) is colored in yellow.

produce lower methylation states of histone H3K9 [53]. Our effort to generate crystals of SUV420H1 and SUV420H2 in complex with a substrate was not successful, but when we superimposed the crystal structures of SUV420H1 and SUV420H2 with the complex structure of GLP-H3K9me2 peptide [56], we found that, like GLP, SUV420H1 and SUV420H2 harbor an aromatic pocket, which presumably would accommodate the target lysine K20 of histone H4, which is confirmed by the recently available mouse SUV420H2 structure in complex with a histone H4K20 peptide (PDB code: 4AU7) (Fig. 4). Interestingly, a glycine G229 in SUV420H1 (G140 in SUV420H2) occupies the position corresponding to Y1124 in GLP, thereby generating extra space next to the methyl-accepting nitrogen. One potential explanation is that a negatively charged residue E225 in SUV420H1 (E136 in SUV420H2) is present at the beginning of the bulge. If E225 took the same conformation like other SET domains, E225 would be pointing toward a hydrophobic core in place of L227, which is not favored (Fig. 4). Although SUV420H1 and SUV420H2 lack an aromatic residue at the G229 position (SUV420H1), they harbor an aromatic residue at W264 (SUV420H1) or W174 (SUV420H2), which occupies a similar position to Y1124 in GLP (Fig. 4). The importance of these active site residues is revealed by our mutagenesis data, showing that G229Y, W264A, S251A, and F281A SUV420H1 mutants are no longer active. The K_m values for monomethylated peptide with G229F (17 ± 1.6 μM) and F311A (8.5 ± 1.7 μM) mutants were higher than that for the wild-type SUV420H1 (1.7 ± 0.3 μM), and their catalytic efficiencies decreased by about 8 and 3 times, respectively (Table 1).

3.5. Trimethylation of H4K20 by SUV420H1 and SUV420H2 is too low to be detectable in vitro

SETD7 is a histone H3K4 monomethylase. Crystal structure of SETD7 in complex with a histone H3K4 peptide reveals that Y245 forms a hydrogen bond with the target lysine through its side chain hydroxyl group, which would not only stabilize the target lysine, but also prevent methylation in this position [4]. Mutating the Y245 to alanine in SETD7 allows SETD7 to trimethylate the histone H3K4 peptide [4]. Consistently, mutating the corresponding residue in G9a (Y1067) could also enable G9a to carry out trimethylation [56], although, in both cases, their enzymatic activity is severely diminished. In the crystal structures of SUV420H1 and SUV420H2, a glycine residue (G229 in SUV420H1) resides in the position corresponding to Y245 in SETD7. Interestingly, a serine residue in SUV420H1 and SUV420H2 points to the same position as Y245 in SETD7 (Fig. 5). Presumably, this serine residue will also form a hydrogen bond with the target lysine through its hydroxyl group. Strikingly, based on the recently released structure of mouse SUV420H2 in complex with a histone H4K20 peptide and S-Adenosyl-L-homocysteine (SAH) (PDB code: 4AU7, Supplementary Fig. 5), which has very high sequence identity (82%) and structure conservation (with a RMSD of 0.58 Å) with its human counterpart, the serine residue S161 in SUV420H2 (S251 in SUV420H1) does form a hydrogen bond with the target lysine through its hydroxyl group (Fig. 5 and Supplementary Fig. 5). Therefore, this structural feature provides a nice explanation for the observation that SUV420H1 and SUV420H2 have virtually no trimethylation activity on H4K20. The importance of this serine residue in catalysis is also reflected by the total loss of activity when S251 is mutated to alanine in SUV420H1 (Table 1). Similarly, mutating the corresponding residue in SETD8 (Y245) also abolished its activity [57]. Taken together, the conserved serine residue in SUV420H1 and SUV420H2 plays two roles: (1) It aligns the target lysine for mono- and di-methylation. (2) It forms a hydrogen bond with the target lysine and controls the methylation product.

In summary, in this work, we presented the high-resolution crystal structures of SUV420H1 and SUV420H2 in complex with

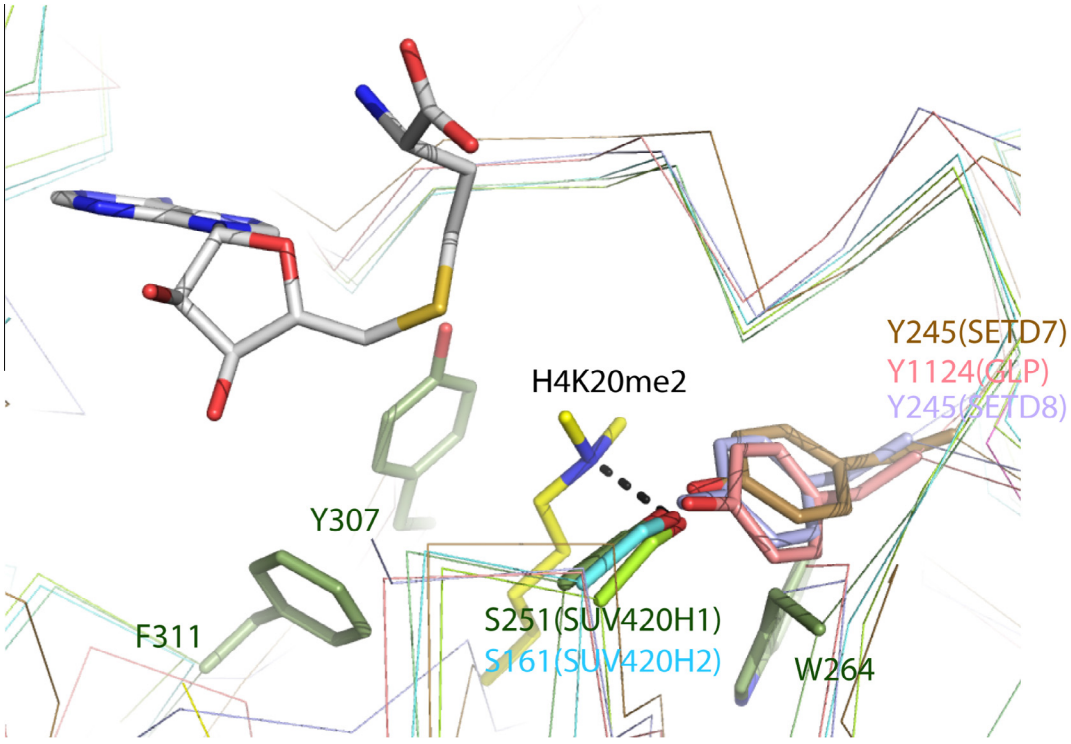

Fig. 5. Superposition of the active sites of SUV420H1 (green), SUV420H2 (human, cyan), SUV420H2 (mouse, limon), GLP (pink, PDB 2RFI), SETD7 (brown, PDB 1O9S) and SETD8 (purple, PDB 1ZKK). S251 in SUV420H1 (S161 in SUV420H2) points to the same position as Y245 in SETD7, Y245 in SETD8 and Y1124 in GLP (or Y1067 in G9a). For clarification, only the cofactor SAH and the histone peptide from the SUV420H2 structure (PDB 4AU7) are shown.

SAM, and characterized their substrate selectivity and product specificity. Our crystal structures and mutagenesis data provide insight to why both SUV420H1 and SUV420H2 are less active with unmodified histone H4K20 peptide (~3 times lower catalytic efficiency) and show no trimethylation activity. By structural comparison, it is also revealed that SUV420H1 and SUV420H2 have unique Zn-binding post-SET and N-terminal domains. Although SUV420H1 and SUV420H2 have similar substrate specificities and potentially could be redundant, they may have different division of labor in the cell, or be under different expression regulations. Since SETD8 is more than 250-fold more active than SUV420H1 and SUV420H2, monomethylation of H4K20 by SETD8 may significantly contribute to the functionality of these two proteins in cells, such as involvement of SUV420H1 in the DNA repair pathway.

Acknowledgments

The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Boehringer Ingelheim, the Canada Foundation for Innovation, the Canadian Institutes for Health Research, Genome Canada through the Ontario Genomics Institute [OGI-055], GlaxoSmithKline, Janssen, Lilly Canada, the Novartis Research Foundation, the Ontario Ministry of Economic Development and Innovation, Pfizer, Takeda, and the Wellcome Trust [092809/Z/10/Z].

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2013.10.020.

References

[1] Martin, C. and Zhang, Y. (2005) The diverse functions of histone lysine methylation. Nat. Rev. Mol. Cell. Biol. 6, 838–849.
[2] Min, J.R., Feng, Q., Li, Z.Z., Zhang, Y. and Xu, R.M. (2003) Structure of the catalytic domain of human DOT1L, a Non-SET domain nucleosomal histone methyltransferase. Cell 112, 711–723.
[3] Lachner, M. and Jenuwein, T. (2002) The many faces of histone lysine methylation. Curr. Opin. Cell. Biol. 14, 286–298.
[4] Xiao, B., Jing, C., Wilson, J.R., Walker, P.A., Vasisht, N., Kelly, G., Howell, S., Taylor, I.A., Blackburn, G.M. and Gamblin, S.J. (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. Nature 421, 652–656.
[5] Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y. and Allis, C.D. (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol. Cell 12, 1591–1598.
[6] Qiao, Q., Li, Y., Chen, Z., Wang, M., Reinberg, D. and Xu, R.M. (2011) The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. J. Biol. Chem. 286, 8361–8368.
[7] Li, Y., Trojer, P., Xu, C.F., Cheung, P., Kuo, A., Drury 3rd, W.J., Qiao, Q., Neubert, T.A., Xu, R.M., Gozani, O. and Reinberg, D. (2009) The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. J. Biol. Chem. 284, 34283–34295.
[8] Kuo, A.J., Cheung, P., Chen, K., Zee, B.M., Kioi, M., Lauring, J., Xi, Y., Park, B.H., Shi, X., Garcia, B.A., Li, W. and Gozani, O. (2011) NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. Mol. Cell 44, 609–620.
[9] Edwards, C.R., Dang, W. and Berger, S.L. (2011) Histone H4 lysine 20 of *Saccharomyces cerevisiae* Is monomethylated and Functions in subtelomeric silencing. Biochemistry 50, 10473–10483.
[10] Nishioka, K., Rice, J.C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., Lis, J.T., Allis, C.D. and Reinberg, D. (2002) PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. Mol. Cell 9, 1201–1213.
[11] Rice, J.C., Nishioka, K., Sarma, K., Steward, R., Reinberg, D. and Allis, C.D. (2002) Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. Genes Dev. 16, 2225–2230.
[12] Fang, J., Feng, Q., Ketel, C.S., Wang, H., Cao, R., Xia, L., Erdjument-Bromage, H., Tempst, P., Simon, J.A. and Zhang, Y. (2002) Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. Curr. Biol. 12, 1086–1099.
[13] Sanders, S.L., Portoso, M., Mata, J., Bahler, J., Allshire, R.C. and Kouzarides, T. (2004) Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. Cell 119, 603–614.
[14] Yang, H. and Mizzen, C.A. (2009) The multiple facets of histone H4-lysine 20 methylation. Biochem. Cell Biol. 87, 151–161.
[15] Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D. and Jenuwein, T. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. 18, 1251–1262.
[16] Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callen, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I.A., Espejo, A., Bedford, M.T., Nussenzweig, A., Busslinger, M. and Jenuwein, T. (2008) A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes Dev. 22, 2048–2061.
[17] Sakaguchi, A., Karachentsev, D., Seth-Pasricha, M., Druzhinina, M. and Steward, R. (2008) Functional characterization of the Drosophila Hmt4-20/Suv4-20 histone methyltransferase. Genetics 179, 317–322.
[18] Beisel, C., Imhof, A., Greene, J., Kremmer, E. and Sauer, F. (2002) Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature 419, 857–862.
[19] Rayasam, G.V., Wendling, O., Angrand, P.O., Mark, M., Niederreither, K., Song, L., Lerouge, T., Hager, G.L., Chambon, P. and Losson, R. (2003) NSD1 is essential for early post-implantation development and has a catalytically active SET domain. Embo J. 22, 3153–3163.
[20] Pei, H., Zhang, L., Luo, K., Qin, Y., Chesi, M., Fei, F., Bergsagel, P.L., Wang, L., You, Z. and Lou, Z. (2011) MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. Nature 470, 124–128.
[21] Yang, H., Pesavento, J.J., Starnes, T.W., Cryderman, D.E., Wallrath, L.L., Kelleher, N.L. and Mizzen, C.A. (2008) Preferential dimethylation of histone H4 lysine 20 by Suv4-20. J. Biol. Chem. 283, 12085–12092.
[22] Trojer, P., Li, G., Sims 3rd, R.J., Vaquero, A., Kalakonda, N., Boccuni, P., Lee, D., Erdjument-Bromage, H., Tempst, P., Nimer, S.D., Wang, Y.H. and Reinberg, D. (2007) L3MBTL1, a histone-methylation-dependent chromatin lock. Cell 129, 915–928.
[23] Kalakonda, N., Fischle, W., Boccuni, P., Gurvich, N., Hoya-Arias, R., Zhao, X., Miyata, Y., Macgrogan, D., Zhang, J., Sims, J.K., Rice, J.C. and Nimer, S.D. (2008) Histone H4 lysine 20 monomethylation promotes transcriptional repression by L3MBTL1. *Oncogene*. 27, 4293–4304.
[24] Li, Z., Nie, F., Wang, S. and Li, L. (2011) Histone H4 Lys 20 monomethylation by histone methylase SET8 mediates Wnt target gene activation. Proc. Natl. Acad. Sci. USA 108, 3116–3123.
[25] Schotta, G. (2011) H4K20 monomethylation faces the WNT. Proc. Natl. Acad. Sci. USA 108, 3097–3098.
[26] Vakoc, C.R., Sachdeva, M.M., Wang, H. and Blobel, G.A. (2006) Profile of histone lysine methylation across transcribed mammalian chromatin. Mol. Cell. Biol. 26, 9185–9195.
[27] Papp, B. and Muller, J. (2006) Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev. 20, 2041–2054.
[28] Du, L.L., Nakamura, T.M. and Russell, P. (2006) Histone modification-dependent and -independent pathways for recruitment of checkpoint protein Crb2 to double-strand breaks. Genes Dev. 20, 1583–1596.
[29] Greeson, N.T., Sengupta, R., Arida, A.R., Jenuwein, T. and Sanders, S.L. (2008) Di-methyl H4 lysine 20 targets the checkpoint protein Crb2 to sites of DNA damage. J. Biol. Chem. 283, 33168–33174.
[30] Botuyan, M.V., Lee, J., Ward, I.M., Kim, J.E., Thompson, J.R., Chen, J. and Mer, G. (2006) Structural basis for the methylation state-specific recognition of histone H4–K20 by 53BP1 and Crb2 in DNA repair. Cell 127, 1361–1373.
[31] Li, H., Fischle, W., Wang, W., Duncan, E.M., Liang, L., Murakami-Ishibe, S., Allis, C.D. and Patel, D.J. (2007) Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger. Mol. Cell 28, 677–691.
[32] Min, J., Allali-Hassani, A., Nady, N., Qi, C., Ouyang, H., Liu, Y., MacKenzie, F., Vedadi, M. and Arrowsmith, C.H. (2007) L3MBTL1 recognition of mono- and dimethylated histones. Nat. Struct. Mol. Biol. 14, 1229–1230.
[33] Adams-Cioaba, M.A. and Min, J.R. (2009) Structure and function of histone methylation binding proteins. Biochem. Cell Biol.-Biochim. Biol. Cell. 87, 93–105.
[34] Guo, Y., Nady, N., Qi, C., Allali-Hassani, A., Zhu, H., Pan, P., Adams-Cioaba, M.A., Amaya, M.F., Dong, A., Vedadi, M., Schapira, M., Read, R.J., Arrowsmith, C.H. and Min, J. (2009) Methylation-state-specific recognition of histones by the MBT repeat protein L3MBTL2. Nucleic Acids Res. 37, 2204–2210.
[35] Eryilmaz, J., Pan, P., Amaya, M.F., Allali-Hassani, A., Dong, A.P., Adams-Cioaba, M.A., MacKenzie, F., Vedadi, M. and Min, J.R. (2009) Structural Studies of a Four-MBT Repeat Protein MBTD1. PLoS One. 4, e7274.
[36] Sakaguchi, A. and Steward, R. (2007) Aberrant monomethylation of histone H4 lysine 20 activates the DNA damage checkpoint in Drosophila melanogaster. J. Cell Biol. 176, 155–162.
[37] Schuetz, A., Allali-Hassani, A., Martin, F., Loppnau, P., Vedadi, M., Bochkarev, A., Plotnikov, A.N., Arrowsmith, C.H. and Min, J. (2006) Structural basis for molecular recognition and presentation of histone H3 by WDR5. EMBO J. 25, 4245–4252.
[38] Otwinowski, Z. and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Meth. Enzymol.* 276, 307–326.

[39] Schneider, T.R. and Sheldrick, G.M. (2002) Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. 58, 1772–1779.
[40] Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. and Paciorek, W. (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Crystallogr. D Biol. Crystallogr. 59, 2023–2030.
[41] Perrakis, A., Harkiolaki, M., Wilson, K.S. and Lamzin, V.S. (2001) ARP/wARP and molecular replacement. Acta Crystallogr. D Biol. Crystallogr. 57, 1445–1450.
[42] Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.
[43] Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 53, 240–255.
[44] Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G.N., Cohen, S., Perrakis, A. and Noble, M. (2004) Developments in the CCP4 molecular-graphics project. Acta Crystallogr. D Biol. Crystallogr. 60, 2288–2294.
[45] Winn, M.D., Isupov, M.N. and Murshudov, G.N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. D Biol. Crystallogr. 57, 122–133.
[46] Vagin, A. and Jones, T.A. (1997) MOLREP: an Automated Program for Molecular Replacement. J. Appl. Cryst. 30, 1022–1025.
[47] Yu, W., Chory, E.J., Wernimont, A.K., Tempel, W., Scopton, A., Federation, A., Marineau, J.J., Qi, J., Barsyte-Lovejoy, D., Yi, J., Marcellus, R., Iacob, R.E., Engen, J.R., Griffin, C., Aman, A., Wienholds, E., Li, F., Pineda, J., Estiu, G., Shatseva, T., Hajian, T., Al-Awar, R., Dick, J.E., Vedadi, M., Brown, P.J., Arrowsmith, C.H., Bradner, J.E. and Schapira, M. (2012) Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors. Nat. Commun. 3, 1288.
[48] Tsang, L.W., Hu, N. and Underhill, D.A. (2010) Comparative analyses of SUV420H1 isoforms and SUV420H2 reveal differences in their cellular localization and effects on myogenic differentiation. PLoS One 5, e14447.
[49] Pesavento, J.J., Yang, H., Kelleher, N.L. and Mizzen, C.A. (2008) Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. Mol. Cell. Biol. 28, 468–486.
[50] Margueron, R., Trojer, P. and Reinberg, D. (2005) The key to development: interpreting the histone code? Curr. Opin. Genet. Dev. 15, 163–176.
[51] Sarg, B., Koutzamani, E., Helliger, W., Rundquist, I. and Lindner, H.H. (2002) Postsynthetic trimethylation of histone H4 at lysine 20 in mammalian tissues is associated with aging. J. Biol. Chem. 277, 39195–39201.
[52] Dillon, S.C., Zhang, X., Trievel, R.C. and Cheng, X. (2005) The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol. 6, 227.
[53] Zhang, X., Yang, Z., Khan, S.I., Horton, J.R., Tamaru, H., Selker, E.U. and Cheng, X. (2003) Structural basis for the product specificity of histone lysine methyltransferases. Mol. Cell 12, 177–185.
[54] Das, A.K., Helps, N.R., Cohen, P.T. and Barford, D. (1996) Crystal structure of the protein serine/threonine phosphatase 2C at 2.0 A resolution. Embo J. 15, 6798–6809.
[55] Nandakumar, J., Shuman, S. and Lima, C.D. (2006) RNA ligase structures reveal the basis for RNA specificity and conformational changes that drive ligation forward. Cell 127, 71–84.
[56] Wu, H., Min, J., Lunin, V.V., Antoshenko, T., Dombrovski, L., Zeng, H., Allali-Hassani, A., Campagna-Slater, V., Vedadi, M., Arrowsmith, C.H., Plotnikov, A.N. and Schapira, M. (2010) Structural biology of human H3K9 methyltransferases. PloS One 5, e8570.
[57] Couture, J.F., Collazo, E., Brunzelle, J.S. and Trievel, R.C. (2005) Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. Genes Dev. 19, 1455–1465.